

WHAT IS CLAIMED IS:

1 1. A method for operating a network of point of presence servers .
2 sharing a hostname, the method comprises:
3 receiving a request from a user for a web page at a first web address, the first
4 web address including the hostname;
5 determining traffic loads of a plurality of customer web servers, each of the
6 customer web servers storing the web page;
7 determining a customer web server from the plurality of customer web servers
8 that is appropriate for the request, the customer web server having a traffic load lower than
9 traffic loads of remaining customer web servers from the plurality of customer web servers;
10 determining an IP address of the customer web server;
11 directing the request from the user to the customer web server; thereafter
12 receiving a request from the user for static content on the web page at a
13 second web address, the second web address including the hostname;
14 determining service metrics of point of presence servers in the network of
15 point of presence servers;
16 determining the point of presence server from the network of point of
17 presence servers that is appropriate for the request, the point of presence server having service
18 metrics better than service metrics of remaining point of presence servers from the network of
19 point of presence servers;
20 retrieving the static content from the point of presence server; and
21 providing the static content to the user.

1 2. The method of claim 1 further comprising determining load of
2 point of presence servers in the network of point of presence servers;
3 wherein determining the point of presence server from the network of point of
4 presence servers that is appropriate for the request, the point of presence server having a
5 latency and a load lower than latency or load of the remaining point of presence servers from
6 the network of point of presence servers.

1 3. The method of claim 1 further comprising:
2 determining whether the point of presence server includes the static content;
3 determining a web server that includes the static content when the point of
4 presence server does not include the static content;

retrieving data from another customer web server from the plurality of customer web servers when the point of presence server does not include the data; and storing the data within the point of presence server.

12. The method of claim 11 wherein retrieving data from the other customer web server comprises:

determining the other customer web server from the plurality of customer web servers, the other customer web server having load measurements lower than load measurements of remaining customer web servers from the plurality of customer web servers; and

retrieving the data from the other customer web server.

13. The method of claim 8 further comprising:

receiving a first request from a second client DNS server to resolve a third domain name, the second client DNS server receiving a request from a second user of a second web page address that includes the third domain name;

determining load measurements of a plurality of second customer web servers, each of the second customer web servers addressable by the third domain name, and each of the second customer web servers storing data configured to service the request from the second user;

determining a second customer web server from the plurality of second customer web servers, the second customer web server having load measurements lower than load measurements of other second customer web servers from the plurality of second customer web servers;

determining an IP address of the second customer web server; and providing the IP address of the second customer web server to the second client DNS server.

14. The method of claim 13 further comprising:

receiving a second request from the second client DNS server to resolve the second domain name, the second client DNS server receiving a request from the second user of a second uniform resource locator that includes the second domain name;

retrieving a second set of data from the point of presence server in response to the second uniform resource locator; and

providing the second set of data to the user.

15. A method for a network of point of presence servers comprises:

receiving at a first point of presence server a first request from a first client DNS server to resolve a first domain name, the first request from the first client DNS server determined in response to a first uniform resource locator entered by a first user, the first uniform resource locator comprising the first domain name;

receiving at a second point of presence server a first request from a second client DNS server to resolve the first domain name, the first request from the second client DNS server determined in response to the first uniform resource locator entered by a second user, the first uniform resource locator comprising the first domain name;

determining at the first point of presence server traffic measurements of a plurality of customer web servers, each of the customer web servers addressable by the first domain name, and each of the customer web servers storing data associated with the first uniform resource locator;

determining at the first point of presence server a first customer web server from the plurality of customer web servers, the first customer web server having traffic measurements lower than load measurements of other customer web servers from the plurality of customer web servers;

determining at the second point of presence server the first customer web server from the plurality of customer web servers, the first customer web server having traffic measurements lower than load measurements of other customer web servers from the plurality of customer web servers;

determining at the first point of presence server an IP address of the first customer web server;

determining at the second point of presence server an IP address of the first customer web server;

providing from the first point of presence server the IP address of the first customer web server to the first client DNS server and to the second client DNS server;

providing from the second point of presence server the IP address of the first customer web server to the first client DNS server and to the second client DNS server;

thereafter

receiving at the first point of presence server a second request from the first client DNS server to resolve a second domain name, the second request from the first client DNS server determined in response to a second uniform resource locator comprising the second domain name, the second uniform resource locator from the first customer web server;

receiving at the second point of presence server a second request from the second client DNS server to resolve the second domain name, the second request from the second client DNS server determined in response to the second uniform resource locator comprising the second domain name, the second uniform resource locator from the first customer web server;

determining at the first point of presence server performance metric measurement between the first point of presence server and other point of presence servers in the network of point of presence servers, each of the point of presence servers addressable by the second domain name;

determining at the second point of presence server performance metric measurement between the second point of presence server and other point of presence servers in the network of point of presence servers;

determining at the first point of presence server performance metric measurement between the second point of presence server and other point of presence servers in the network of point of presence servers;

determining at the first point of presence server a third point of presence server from the network of point of presence servers, the third point of presence server having performance metric measurements lower than performance metric measurements with regards to the first point of presence server;

determining at the second point of presence server a fourth point of presence server from the network of point of presence servers, the fourth point of presence server having performance metric measurements lower than performance metric measurements with regards to the second point of presence server;

retrieving data from a third point of presence server;

retrieving the data from the fourth point of presence server;

providing the data to the first user; and

providing the data to the second user.

16. The method of claim 15 wherein the traffic measurements comprise latency measurements.

17. The method of claim 15 wherein the performance metric measurements comprise CPU load measurements, HTTP response measurements, FTP response measurements.

